

# FITTING A STOCHASTIC CLUSTER MODEL (GNSRP(2)) TO HOURLY PRECIPITATION AT DAROCA OBSERVATORY

José Antonio LÓPEZ DÍAZ  
Agencia Estatal de Meteorología  
jlopezd@aemet.es

## RESUMEN

Los procesos estocásticos del tipo Generalised Neyman-Scott Rectangular Pulses con dos tipos de célula (*GNSRP(2)*) poseen, según la literatura, la suficiente complejidad como para simular realísticamente la precipitación horaria. Consisten en un proceso de Poisson para modelar el origen de las tormentas, cada una de las cuales produce un clúster con un número aleatorio de células (aquí modelado como una variable aleatoria geométrica). Cada célula tiene, independientemente, una probabilidad fija de pertenecer a uno de los dos tipos de células. Cada tipo de célula se modela como un pulso rectangular de lluvia, esto es, una duración aleatoria con una intensidad constante. Tanto la duración aleatoria como la intensidad se modelan como variables aleatorias exponenciales, igual que el retraso entre el origen de la tormenta y el inicio de cada célula. Estos procesos pueden, en teoría, simular los dos tipos principales de precipitación en nuestras latitudes: la convectiva y la frontal. En el ajuste a los datos de precipitación horaria de Daroca, se muestra en este trabajo que el *GNSRP(1)*, que sólo tiene un tipo de célula, proporciona un ajuste aceptable para algunos meses del año. Un inconveniente de este tipo de proceso estocástico es el gran número de parámetros que tienen que ser simulados, 8 para el *GNSRP(2)*. Para este fin se utiliza un algoritmo numérico.

**Palabras clave:** Proceso Clúster, Precipitación Horaria, Proceso Estocástico.

## ABSTRACT

The stochastic processes of the type Generalised Neyman-Scott Rectangular Pulses with two types of cell (*GNSRP(2)*) have been shown in the literature to be of sufficient complexity to realistically simulate the hourly precipitation. They consist of a Poisson process modelling the origin of storms with each storm producing a cluster with a random number of cells (here modelled as a geometric random variable). Each cell independently has a fixed probability of belonging to either of two types of cells. Each type of cell is modelled as a rectangular pulse of rain, i.e., a random duration of time with a constant random intensity. Both the random duration and random intensity are modelled as exponential random variables, as well as the random time delay from the origin of the storm for the onset of each cell. In theory these process can simulate the two main types of precipitation in our latitudes, namely the convective and the frontal precipitation. For the case of hourly precipitation data at Daroca observatory, it is shown in this work that the *GNSRP(1)*, i.e., with only one type of cell, gives a good enough fit for some months of the year. One drawback of this type of stochastic model is the large number of parameters, 8

for the  $GNSRP(2)$ , that need to be estimated. For this purpose a numerical optimization algorithm is used.

**Key words:** Cluster Process, Hourly Precipitation, Stochastic Process.

## 1. INTRODUCCIÓN

En los últimos 25 años los hidrólogos han venido usando la teoría de procesos puntuales (Cox, 1980) para desarrollar modelos capaces de generar series sintéticas de precipitación. En teoría, la salida de estos modelos debe tener características estadísticas similares a las de los datos de precipitación usados en el ajuste del modelo. El interés primordial en estos modelos radica en su capacidad para desagregar los intervalos temporales hacia mayores resoluciones de lo que es posible con un pluviómetro de una resolución dada. Por ejemplo, con un modelo ajustado a partir de dato diario de precipitación, el modelo puede producir una serie sintética de precipitación horaria.

La mayoría de estos modelos estocásticos de precipitación se basan en el proceso de Poisson, que se describe brevemente en el apartado 3. De los modelos basados en el proceso de Poisson quizá los más referenciados sean los procesos de Poisson cluster con pulsos rectangulares, descritos por primera vez en Rodríguez-Iturbe *et al* (1987). Hay dos versiones de este modelo, una basada en el proceso de Neyman-Scott y la otra basada en el proceso de Barlett-Lewis. El proceso de Neyman-Scott asume que los desplazamientos de cada célula respecto a su tormenta origen dada son independientes estadísticamente y tienen distribución exponencial, este tipo es el utilizado en este artículo. En cambio, en el proceso de Barlett-Lewis los intervalos entre células sucesivas están determinados por un proceso de Poisson. Después de un tiempo, que está distribuido exponencialmente, el proceso de Poisson de generación de células termina. Una referencia reciente con una aplicación de una superposición de un proceso de Neyman-Scott de pulsos rectangulares y de un proceso de Poisson de ruido blanco se encuentra en Morrissey (2009).

## 2. DATOS

Los datos utilizados para este estudio han sido los de precipitación horaria en el observatorio de Daroca (Zaragoza) procedentes del banco de datos climatológico de AEMET. Este observatorio se ha seleccionado para este estudio por la longitud de su serie de precipitación horaria. Los datos horarios de precipitación comienzan en 1982 en Daroca. Se han seleccionado solo aquellos meses que no presentaban ninguna laguna en los datos horarios. Han resultado para los sucesivos meses del año un número de meses completos de 25, 27, 30, 26, 30, 29, 29, 29, 29, 29, 27 y 24.

## 3. METODOLOGÍA

### 3.1. El proceso estocástico $GNSRP(2)$

Se define a continuación el proceso puntual de clúster generalizado de Neyman-Scott de pulsos rectangulares con dos tipos de célula  $GNSRP(2)$  (Cowpertwait, 1994). Suponemos que los orígenes de las tormentas ocurren según un proceso de Poisson de tasa (por unidad de tiempo)  $\lambda$  y que un número aleatorio  $C$  de orígenes de células se asocia con cada origen de tormenta. En este trabajo la distribución de  $C$  se ha supuesto geométrica con parámetro  $v$ , por tanto con media  $v^{-1}$ . Recordemos

que para un proceso de Poisson de tasa  $\lambda$  en cada intervalo temporal infinitesimal  $\delta t$  la probabilidad de que caiga un origen de tormenta es  $\lambda \delta t$ , y los sucesos correspondientes a intervalos distintos son independientes. Los orígenes de células están retrasados de sus orígenes de tormenta por distancias que son variables aleatorias exponenciales con parámetro  $\beta$ . Cada célula independientemente es clasificada como de tipo *I* o *II*, con probabilidades respectivas  $\alpha_1$  y  $\alpha_2 = 1 - \alpha_1$ . Un pulso rectangular de lluvia es asociado independientemente con cada origen de célula. La duración del pulso es una variable exponencial independiente de parámetro  $\eta_i$  para el tipo  $i$  de célula,  $i = 1, 2$ . La intensidad del pulso, notada  $X_i$ , es también una variable exponencial independiente de parámetro  $\iota_i$  según el tipo de célula,  $i = 1, 2$ .

Denotemos con  $Y(t)$  el total de intensidad de precipitación en el tiempo  $t$  dado por el modelo GNSRP(2), y sea  $X_{t-u}(u)$  la intensidad de precipitación en el tiempo  $t$  debida a una célula con origen en  $t-u$ . Entonces  $X_{t-u}(u)$  será igual a  $X_1$  con probabilidad  $\alpha_1 \text{Exp}(-\eta_1 u)$ , a  $X_2$  con probabilidad  $(1 - \alpha_1) \text{Exp}(-\eta_2 u)$ , y 0 en otro caso.  $Y(t)$  es la suma de todas la células activas en el tiempo  $t$ , o sea,

$$(1) \quad Y(t) = \int_{u=0}^{\infty} X_{t-u}(u) dN(t-u)$$

en donde  $dN(t-u) = 1$  si hay un origen de célula en  $t-u$ , y 0 en caso contrario.

Para estimar los parámetros del modelo necesitamos sus propiedades estadísticas agregadas, pues los datos de que disponemos están en forma agregada, como totales horarios o diarios. Notemos por  $Y_i^{(h)}$  la precipitación agregada en el  $i$ -ésimo intervalo temporal de longitud  $h$ , es decir

$$(2) \quad Y_i^{(h)} = \int_{(i-1)h}^{ih} Y(t) dt$$

Por tanto la serie  $\{Y_i^{(h)} : i = 1, 2, \dots\}$  es una serie temporal de precipitación al nivel  $h$  de agregación. Las propiedades estadísticas de segundo orden del proceso agregado vienen dadas por (Cowpertwait, 1994):

$$(3) \quad \begin{aligned} E(Y_i^{(h)}) &= h\lambda v^{-1} \sum_{j=1}^2 \alpha_j / \iota_j \eta_j \\ \text{Var}\{Y_i^{(h)}\} &= \sum_{j=1}^2 \left[ \frac{2C_j}{\eta_j^2} (h\eta_j + \text{Exp}(-h\eta_j) - 1) \right] + \frac{2C_\beta}{\beta^2} (h\beta + \text{Exp}(-h\beta) - 1) \end{aligned}$$

y para  $k \geq 1$ ,

$$(4) \quad \text{Cov}\{Y_i^{(h)}, Y_{i+k}^{(h)}\} = \sum_{j=1}^2 \left[ \frac{2C_j}{\eta_j^2} \text{Exp}[-\eta_j(k-1)h] (1 - \text{Exp}(-h\eta_j))^2 \right] + \frac{2C_\beta}{\beta^2} \text{Exp}[\beta(k-1)h] (1 - \text{Exp}(-h\beta))^2$$

En estas expresiones las definiciones para  $C_j$  y  $C_\beta$  son:

$$C_j = \frac{\lambda v^{-1} \alpha_j E(X_j^2)}{\eta_j} + \frac{\lambda \beta^2 \iota_j^{-1} \alpha_j E(C^2 - C)}{\beta^2 - \eta_j^2} - \sum_{i=1}^2 \frac{\iota_i^{-1} \alpha_i}{\eta_i + \eta_j}, \quad j = 1, 2,$$

$$\text{y} \quad C_\beta = -\frac{1}{2} \lambda \beta E(C^2 - C) - \sum_{i=1}^2 \sum_{j=1}^2 \{\iota_i^{-1} \iota_j^{-1} \alpha_i \alpha_j / [(\beta - \eta_j)(\beta + \eta_i)]\}$$

Para calcular la probabilidad de que no llueva en un intervalo arbitrario de longitud  $h$  partimos de la probabilidad  $p_h(t)$  de que no llueva en el intervalo  $(t, t+h)$  debido a una tormenta con origen en  $t=0$ :

$$p_h(t) = \frac{v [Exp[-\beta(t+h)] + \omega(1 - Exp(-\beta t))]}{1 - (1-v) [Exp[-\beta(t+h)] + \omega(1 - Exp(-\beta t))]},$$

$$\text{siendo } \omega = \sum_{i=1}^2 \alpha_i \{ -\beta [Exp(-\beta t) - Exp(-\eta_i t)] / [(\eta_i - \beta)(1 - Exp(-\beta t))] \}$$

La probabilidad  $\varphi(h)$  de que no llueva en un intervalo arbitrario de longitud  $h$  está dada por

$$\begin{aligned} \varphi(h) &= Exp \left\{ -\lambda \int_0^\infty [1 - p_h(t)] dt \right\} Exp \left\{ -\lambda \int_0^h [1 - p_t(0)] dt \right\} = \\ (5) \quad &= Exp \left\{ -\lambda \int_0^\infty [1 - p_h(t)] dt \right\} Exp \left\{ -\lambda h + \frac{\lambda v}{\beta(1-v)} \ln [v^{-1} + (1-v^{-1}) Exp(-\beta h)] \right\} \end{aligned}$$

De  $\varphi(h)$  se puede derivar la probabilidad  $\varphi_{ww}(h)$  de que un intervalo arbitrario de longitud  $h$  sea húmedo (con precipitación no nula) condicionado a que el anterior intervalo de longitud  $h$  también sea húmedo:

$$(6) \quad \phi_{ww}(h) = [1 - 2\varphi(h) + \varphi(2h)] / [1 - \varphi(h)]$$

### 3.2 Función de coste para el ajuste a los datos del proceso *GNSRP(2)*

Para cada mes del año se ha procedido a ajustar un proceso *GNSRP(2)* a los datos de precipitación horaria de Daroca. Para especificar el proceso *GNSRP(2)* son necesarios 8 parámetros:  $\lambda, v, \beta, \alpha_1, \eta_1, \iota_1, \eta_2, \iota_2$ . Para medir el grado de disparidad entre el proceso estimado *GNSRP(2)* y los datos de precipitación se ha utilizado una suma de cuadrados de errores relativos de estadísticos del proceso agregado, de la forma:

$$(7) \quad SS = \sum_{i \in H} \sum_{g_i \in G} \left\{ \left( 1 - \frac{g_i(h)}{\hat{g}_i(h)} \right)^2 + \left( 1 - \frac{\hat{g}_i(h)}{g_i(h)} \right)^2 \right\}$$

donde  $G$  es un conjunto de propiedades estadísticas agregadas para el modelo *GNSRP(2)*,  $\hat{g}_i$  denota el estimador muestral (a partir de los datos) del estadístico teórico del proceso  $g_i$  y  $H$  es un conjunto de niveles de agregación temporal.

Existen claramente muchas posibilidades para escoger  $G$  y  $H$ . En este trabajo se eligieron las siguientes: media de 1 h (ver (3)), para las agregaciones de 1, 3, 6, 12 y 24 h las varianzas (ver (3)), la proporción de intervalos secos (ver (5)) y la probabilidad de transición de húmedo a húmedo (ver (6)).

### 3.3. El cálculo de $\varphi(h)$

El cálculo de  $\varphi(h)$  dado por (5) presenta problemas en la primera exponencial  $Exp \left\{ -\lambda \int_0^\infty [1 - p_h(t)] dt \right\}$

pues pequeños errores numérico en la integral conducen a errores importantes en la exponencial. La cuestión estriba en que, de la definición de  $p_h(t)$ , se ve que  $p_h(t) \rightarrow 1$  cuando  $t \rightarrow \infty$  y por tanto

$(1 - p_h(t)) \rightarrow 0$  cuando  $t \rightarrow \infty$ . Pequeños errores en  $p_h(t)$  hacen que la exponencial de la integral oscile. Así que fue necesaria la búsqueda de una cota superior del error que se comete al sustituir en la integral el límite superior  $\infty$  por un valor grande pero finito.

Analizando la estructura de  $p_h(t)$  se llegó a la siguiente fórmula para una cota superior de dicha integral, en que  $\eta = \min(\eta_1, \eta_2)$ :

$$-v^{-1} \left( -\frac{e^{-t\beta}}{\beta - \eta} + \frac{e^{-(h+t)\beta}}{\beta - \eta} - \frac{e^{-(h+t)\beta}\eta}{\beta(\beta - \eta)} + \frac{e^{-t\eta}\beta}{\eta(-\beta + \eta)} + \frac{e^{-t(\beta+\eta)}\beta}{(\beta - \eta)(\beta + \eta)} - \frac{e^{-2t\beta}\eta}{2\beta^2 - 2\beta\eta} + \frac{2e^{-t\beta}\eta}{\beta^2 - \beta\eta} \right)$$

### 3.4. Algoritmo numérico de optimización

Para minimizar (7) respecto a los 8 parámetros del proceso *GNSRP(2)* desconocidos se ha utilizado el algoritmo de Nelder-Mead. Este método usa el concepto de simplex, que es una generalización de una línea en dos dimensiones, un triángulo en tres o un tetraedro en cuatro dimensiones, y es un polítopo de  $N+1$  vértices en  $N$  dimensiones. El algoritmo de Nelder-Mead genera una nueva posición de prueba a partir del comportamiento de la función objetivo en los vértices del simplex. Por ejemplo se puede reemplazar el peor punto del simplex por medio de una reflexión respecto al centroide de los puntos restantes del simplex. Si este punto es mejor que los anteriores se alarga exponencialmente a lo largo de esta línea. En cambio si el nuevo punto no mejora sustancialmente los anteriores, entonces probablemente estemos en una zona de valle, y lo que hacemos es encoger el simplex hacia un nuevo punto.

Realizando diversas pruebas de minimización de la función de coste (7) para distintos meses partiendo de distintas condiciones iniciales se vio que el mínimo obtenido variaba, por lo que se ha utilizado un procedimiento de búsqueda aleatoria previa de condiciones iniciales favorables antes de lanzar el algoritmo numérico de optimización. Este proceso se describe en el próximo apartado.

## 4. AJUSTE A LOS DATOS

Para el ajuste del proceso *GNSRP(2)* a los datos de precipitación horaria de cada mes se han seguido varios pasos, incluyendo el ajuste del modelo más sencillo *GNSRP(1)*, con un solo tipo de célula, que en algunos meses ha resultado ser preferible al *GNSRP(2)*.

### 4.1. Ajuste inicial de modelos *GNSRP(1)* y *GNSRP(2)*

Para cada mes del año se ha procedido a ajustar los modelos *GNSRP(1)* y *GNSRP(2)* a los datos horarios con una búsqueda previa aleatoria de condiciones iniciales favorables y posterior búsqueda del óptimo de la función de coste con el algoritmo de Nelder-Mead. Los términos empíricos en la función de coste (7) se han calculado a partir de los datos horarios de cada mes concatenando años sucesivos. Para el proceso *GNSRP(2)* se ha evaluado la función de coste (7) 1000 veces variando los parámetros del proceso aleatoriamente según las condiciones:  $\lambda$  correspondiente a un número medio aleatorio de tormentas en un mes entre 30 y 1; el número medio de células por tormenta, dado por  $v$ , entre 1 y 20;  $\beta$  derivado de un retraso medio de células respecto a tormenta entre 1 y 72 horas; la probabilidad de los dos tipos de células dada por  $\alpha_1$  entre 1% y 99%; la escala temporal de la duración de células dada por  $\eta_i$  como la de  $\beta$ ; para las dos intensidades de cada tipo de células se ha usado la ligadura dada por la precipitación media

horaria, igualando el valor teórico dado por (3) con el empírico con un factor de variación aleatorio entre el 80% y el 120%, para derivar  $\iota_2$  de  $\iota_1$  a partir de otro factor aleatorio. Para el  $GNSRP(1)$  se ha usado un procedimiento similar.

Una vez calculado el valor óptimo de las 1000 simulaciones previas aleatorias, se ha usado esa solución como condición inicial para lanzar el algoritmo numérico de Nelder-Mead.

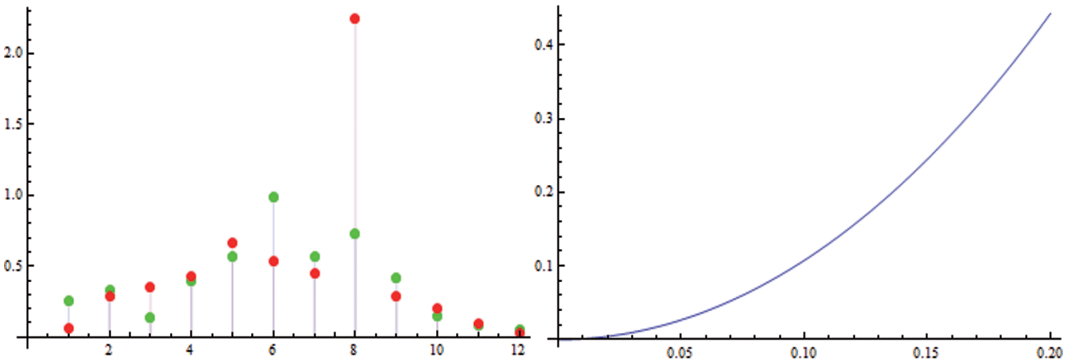


FIG. 1: Izqda: función de coste (7) para cada mes con el ajuste inicial de  $GNSRP(1)$  (verde) y  $GNSRP(2)$  (rojo). Drcha: valor esperado de (7) suponiendo variación de los errores relativos entre  $-t$  y  $t$  con  $t$  (en abscisas)

Con este ajuste inicial de ambos procesos los errores cuadráticos finales obtenidos para cada mes se representan en la fig. 1. Para valorar su orden de magnitud en la parte derecha de la fig.1 se ha dibujado la curva que da el valor esperado de  $16 [(1-X)^2 + (1-X^{-1})^2]$  con  $X$  distribuida uniformemente en  $(1 - t, 1 + t)$ , con  $t$  en abscisas entre 0.0 y 0.2. Esto supone suponer que en cada uno de los 16 sumandos de (7) los cocientes entre magnitudes teóricas y empíricas varían todos entre los márgenes dados. La figura 1 pone de manifiesto que para algunos meses, de forma destacada agosto, el ajuste con más parámetros da peor valor de la función de coste que con menos parámetros, lo cual no es una situación aceptable, ya que el modelo  $GNSRP(1)$  es un caso particular del  $GNSRP(2)$ . Pero esto pone de manifiesto las dificultades encontradas en la práctica con el ajuste de modelos de tantos parámetros.

4.2 Ajuste final de modelos  $GNSRP(1)$  y  $GNSRP(2)$

Por tanto lo que se ha hecho a continuación es proceder a un ajuste de  $GNSRP(2)$  partiendo de las condiciones iniciales dadas por el ajuste a  $GNSRP(1)$  anterior para los meses del 2 al 10. Con esto se ha mejorado notablemente el ajuste final a  $GNSRP(2)$ , obteniéndose los errores cuadráticos de la figura 2.

Nótese en la figura 2 la considerable disminución del error para ajuste a  $GNSRP(2)$  en agosto, abril o marzo. Ahora ya todos los errores del ajuste para el modelo más complejo son inferiores que para el modelo más simple.

Para decidir en cada mes con qué modelo quedarnos se ha tenido en cuenta que el modelo más complejo debe disminuir de forma aceptable la función de coste para ser preferible al modelo más simple. Se ha usado el criterio de considerar aceptable para un mes el modelo  $GNSRP(1)$  si el valor de la función de coste (7) es inferior a 0.1, y si el valor de (7) es inferior a 1.2 veces el valor de (7)

para dos células. En caso contrario se toma el *GNSRP(2)*. Con este criterio la elección para cada mes queda: ene(2), feb(1), mar(2), abr(2), may(2), jun(2), jul(2), ago(2), sep(2), oct(1), nov(1), dic(1).

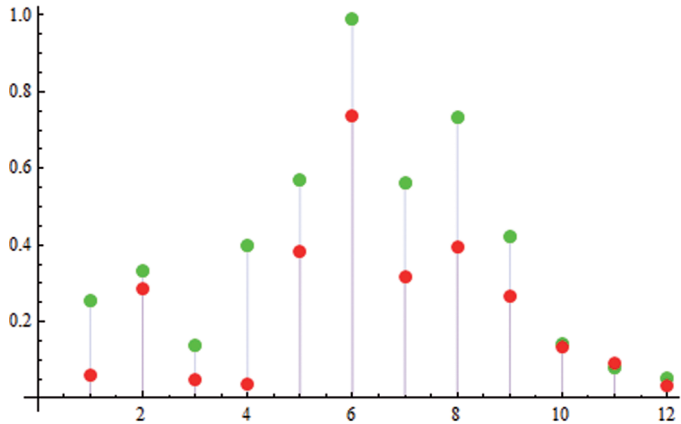


FIG. 2: Función de coste (7) para ajuste de *GNSRP(1)* (verde) y *GNSRP(2)* (rojo) en la segundo fase.

Pero falta un último paso, porque analizando las probabilidades de los dos tipos de células en los meses con modelo *GNSRP(2)* en la lista anterior se ve que marzo y abril dan una probabilidad muy pequeña a unos de los tipos de célula, del orden de un 0.1 %. Por tanto es natural probar si reduciendo el modelo a *GNSRP(1)* eliminando el tipo de célula de baja probabilidad, el valor de la función de coste (7) no se incrementa. Esto se comprobó que era cierto, y por tanto estos dos meses pasaron a *GNSRP(1)*, con lo que la asignación final de modelos a cada mes queda: ene(2), feb(1), mar(1), abr(1), may(2), jun(2), jul(2), ago(2), sep(2), oct(1), nov(1), dic(1). Los errores finales del ajuste con esta selección se dibujan en la figura 3.

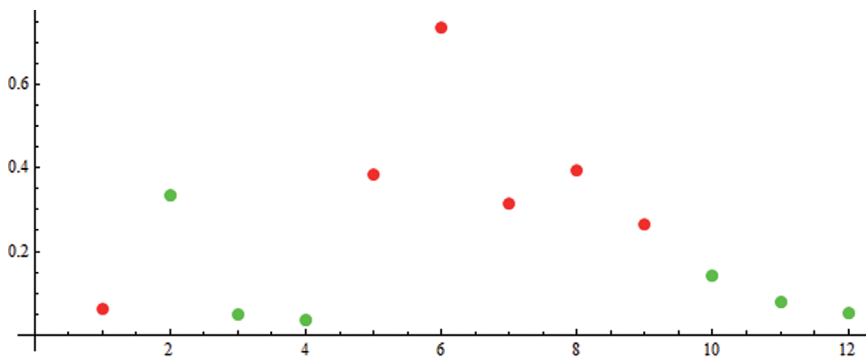


FIG. 3: Función de coste (7) para ajuste de *GNSRP(1)* (verde) y *GNSRP(2)* (rojo) en el ajuste final.

5. SIMULACIONES CON EL PROCESO *GNSRP* AJUSTADO

Al objeto de validar el ajuste de los procesos de tipo *GNSRP* a los datos de cada mes se han simulado 30\*24\*1000 horas de precipitación cada mes del año con proceso *GNSRP(1)* o *GNSRP(2)* ajustado al mes, según el caso. Por tanto para cada mes del año se han generado 1000 meses de

precipitación simulada. Un primer aspecto que podemos comparar es el histograma de precipitación total mensual simulada y empírica. Estos histogramas se comparan en las figuras 4 y 5.

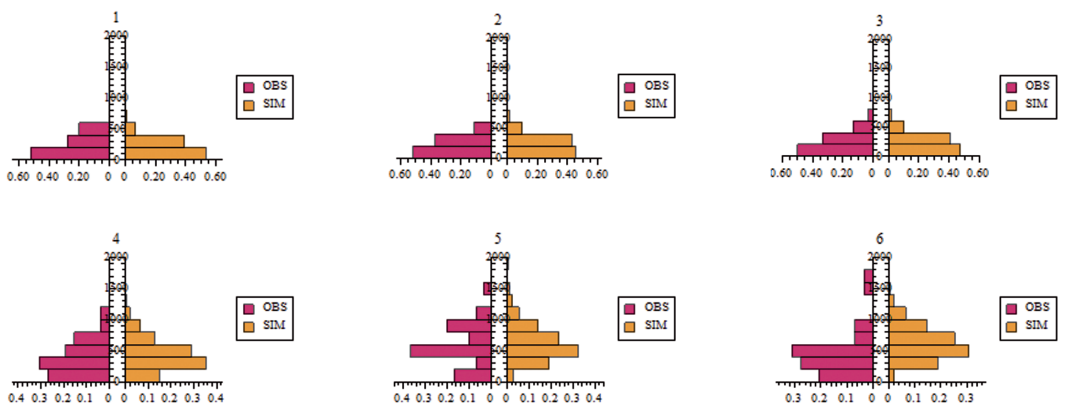


FIG. 4: Histogramas para los 6 primeros meses del años de precipitación total mensual (mm) en cajas de 200 mm (malva observado, naranja simulado).

Una característica que cabe observar en estas comparativas mensuales es que el histograma simulado es más regular, en el sentido de presentar una mono-modalidad (máximo único) más claramente dibujada, con decrecimiento monótono regular a partir del máximo. Este máximo en algunos meses se da en el intervalo de precipitación más próximo al origen, y por tanto el histograma es monótonamente decreciente. Esto sucede en los meses de ene, feb, mar, jul, ago, sep, dic, que son básicamente los invernales y veraniegos. En el resto de los meses el máximo corresponde a un intervalo de precipitación más central, con decrecimiento a ambos lados desde él. En cambio en los histogramas empíricos vemos más irregularidades en algunos meses, como mayo, junio o septiembre. En los meses de mayo y junio esta irregularidad da lugar a que en el intervalo de menor precipitación, de 0 a 200 mm, el histograma empírico tenga mucha más probabilidad que el simulado.

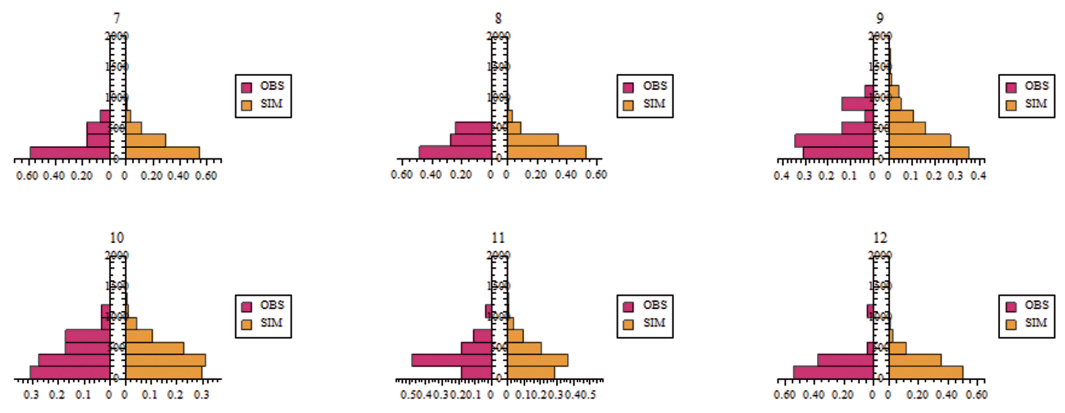


FIG. 5: Histogramas para los 6 últimos meses del años de precipitación total mensual (mm) en cajas de 200 mm (malva observado, naranja simulado).



Por otra parte es claro que el histograma simulado debe tener un aspecto efectivamente regular, al corresponder a un proceso supuesto estacionario, mientras las irregularidades de algunos meses en los histogramas empíricos pueden responder a peculiaridades del muestreo de un número limitado de años.

Para el total anual de precipitación se representa en la figura 6 una comparación de funciones de densidad empírica y ajustada, obtenida aplicando un kernel gaussiano. La densidad simulada es más simétrica que la obtenida a partir de los datos, y tiene una cola izquierda bastante menos larga que la observada. Esta insuficiente representación de los años más secos en el proceso ajustado se podría explicar por el hecho de que el ajuste se ha hecho mes a mes, y por tanto el total anual simulado no puede capturar las correlaciones entre meses que sin duda están presentes en los años más secos (al combinarse los meses del año de forma aleatoria en la simulación). Esto también explicaría la mayor simetría. En cambio la cola derecha, correspondiente a años muy húmedos, sí que encaja muy bien con la observada.

## 6. CONCLUSIONES

En este trabajo se ha descrito el proceso estocástico  $GNSRP(2)$  que ha sido el modelo base utilizado para ajustar a la serie de precipitación horaria de cada mes en Daroca. El proceso de ajuste ha tenido que dividirse en varias etapas, debido fundamentalmente a que la gran cantidad de parámetros del proceso  $GNSRP(2)$  hace que surjan diversos problemas para hallar una solución óptima. El procedimiento que se ha seguido finalmente incluye una exploración previa del espacio paramétrico de 5 dimensiones del  $GNSRP(1)$  (proceso con un solo tipo de células) con 1000 simulaciones, y después selección de la mejor solución para proceder al ajuste del  $GNSRP(2)$  con esa condición inicial.

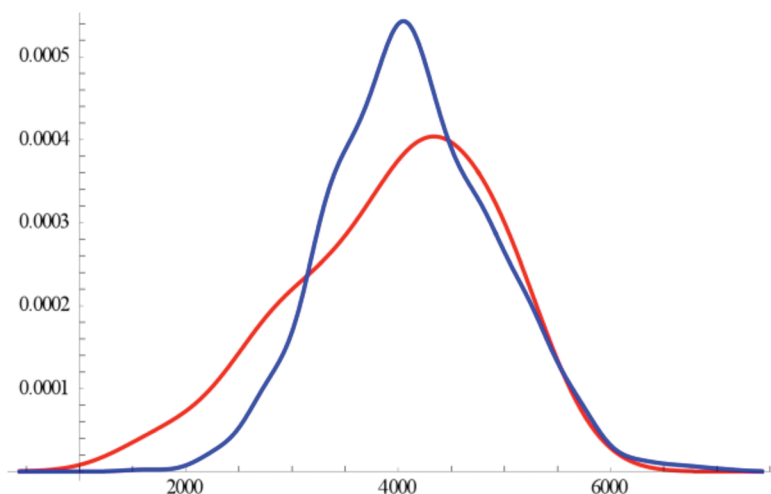


FIG. 6: Funciones de densidad de la precipitación total anual obtenidas aplicando un kernel gaussiano. En rojo, función de densidad de los totales observados de los últimos 30 años en Daroca; en azul, función de densidad a partir del  $GNSRP$  ajustado.

Se ha encontrado que en varios meses del año, en primavera y otoño, el ajuste a un proceso  $GNSRP(1)$  (proceso con un solo tipo de células) ha sido suficiente. La comprobación de los procesos

ajustados a cada mes por medio de los histogramas de precipitación total mensual empírica y obtenida mediante simulación del proceso ajustado muestran en general una buena concordancia. En los meses de mayo y junio es donde hay mayor discrepancia, en el sentido de déficit en la clase de menor precipitación en la simulación. Para el total anual la misma comparación entre histogramas empírico y simulado muestra que la cola de la derecha está muy bien capturada, pero de nuevo para valores bajos de la precipitación total el histograma simulado es deficitario.

## REFERENCIAS

- O'Neill, R. (1971). "Algorithm AS 47: Function Minimization Using a Simplex Procedure". *Journal of the Royal Stat. Soc. Series C (Applied Statistics)*, Vol. 20, No. 3, pp. 338-34.
- Cowpertwait, P.S.P. (1994). "A Generalised Point Process Model for Rainfall". *Proc. R. Soc. Lond. A* 447, 23-27
- Cox, D. R., y Isham, V. (1980). *Point Processes*. Chapman and Hall, 188 pp.
- Morrissey, M. L. (2009). "Superposition of the Neyman-Scott Rectangular Pulses Model and the Poisson White Noise Model for the Representation of Tropical Rain Rates". *J. Of Hydrometeorology*, vol. 10, pp. 395- 411.
- Rodríguez-Iturbe, I.; Cox, D. y Isham, V. (1987). "Some models for rainfall based on stochastic point processes". *Proc. Roy. Soc. London*, A410, 269-288.